



(12) **United States Patent**
Waite et al.

(10) **Patent No.:** **US 7,797,152 B1**
(45) **Date of Patent:** **Sep. 14, 2010**

(54) **METHOD OF DATABASE SEARCHING**

(75) Inventors: **David P. Waite**, Silver Spring, MD (US);
Richard O. Wyckoff, Woodstock, MD (US)

(73) Assignee: **The United States of America as represented by the Director, National Security Agency**, Washington, DC (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 723 days.

(21) Appl. No.: **11/363,299**

(22) Filed: **Feb. 17, 2006**

(51) **Int. Cl.**
G06F 17/27 (2006.01)

(52) **U.S. Cl.** **704/9; 704/10; 707/3; 707/6; 382/187; 382/228**

(58) **Field of Classification Search** 704/9, 704/10; 707/3, 6; 382/187, 228
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,392,363 A * 2/1995 Fujisaki et al. 382/187
5,768,423 A * 6/1998 Aref et al. 382/228
5,835,888 A * 11/1998 Kanevsky et al. 704/9
6,018,708 A 1/2000 Dahan et al.
6,182,039 B1 1/2001 Rigazio et al.

6,415,250 B1 * 7/2002 van den Akker 704/9
6,557,004 B1 4/2003 Ben-Shachar et al.
6,662,180 B1 * 12/2003 Aref et al. 707/6
7,043,494 B1 * 5/2006 Joshi et al. 707/101
2002/0002550 A1 * 1/2002 Berman 707/3
2004/0107089 A1 * 6/2004 Gross et al. 704/10
2005/0108001 A1 * 5/2005 Aarskog 704/10

* cited by examiner

Primary Examiner—Richemond Dorvil

Assistant Examiner—Michael C Colucci

(74) *Attorney, Agent, or Firm*—Jennifer P. Ferragut; Robert D. Morelli

(57) **ABSTRACT**

The present invention is a method of database searching. First, a language is selected and elements received. The system is searched to identify a unit number associated with each element, which is linked to a data unit containing morphological variants of the element. If none are identified, the element is broken into sub-textual units that may contain a prefix, compound-prefix, and/or suffix along with a primary element. A unit number is then obtained for the primary element. If this does not result in a match, the elements may be saved in a database for further linguistic development. A unit number associated with each matched element is then chosen, and the elements contained in the data units linked to the unit numbers are compared to a database index. If an element is associated with multiple unit numbers, this process is repeated until all data units have been compared to the database.

13 Claims, 3 Drawing Sheets

